Characterization and Sequence Verification of Thiolated Deoxyoligonucleotides Used for Microarray Construction

Arthur Van Aerschot and Jef Rozenski
Rega Institute for Medical Research, Faculty of Pharmaceutical Sciences, Katholieke Universiteit Leuven, Leuven, Belgium

During synthesis of thiolated deoxyoligonucleotides, side products can be formed. When used in the fabrication of microchips, the oligonucleotides have to be of high purity. We demonstrated the presence of impurities, which were not failure sequences from the synthesis. These products were identified and characterized using high-performance liquid chromatography (HPLC) and electrospray MS(/MS). The presence of the free thiol group was assessed by acrylamide derivatization. After reaction with acrylamide the correct compounds showed a 71 u mass shift and the major fragment ions could be assigned as 5′ a-base and 3′ w ions, similar as for unmodified DNA. The side products were unaffected by acrylamide, suggesting that the thiol group was modified. The oligonucleotide containing a species with a mass of 32 u higher was identified as 5′ sulfinic acid containing molecule and was found as 45% of the total amount of a DNA 25 mer. Other components appeared to be dithio-linked oxidation products, present about 1 to 5% in a 10 mer and 15 mer deoxyoligonucleotide. The analyses were useful for optimizing the synthesis protocols. (J Am Soc Mass Spectrom 2006, 17, 1396–1400) © 2006 American Society for Mass Spectrometry

Microarrays or DNA chips have become a standard tool for diagnostics and for molecular biology research. They can be applied for the detection of single-nucleotide polymorphism (SNPs) as a genetic marker in the study of human diseases [1, 2].

Manufacturing of DNA chips can be done by either direct synthesis of DNA on a substrate, or by coupling presynthesized DNA to a microchip. For the latter method, generally aminated or thiolated oligonucleotides are used [3–5]. The key problems to overcome for the preparation of DNA array based sensors or chips are the attachment of functional groups to the solid surface, the synthesis of oligonucleotides containing reactive groups, the spacial distribution of the DNA oligomers, and the detection of hybridization with complementary strands. For the fabrication of DNA arrays, it is important the oligonucleotides are be free of contaminations, especially structurally related molecules. Our work consisted in the synthesis and analysis of 5′ thiol containing deoxyoligonucleotides. Several methods are available for the synthesis of thiolated oligonucleotides [6–8]. The approach using trityl protected 5′ thio building blocks has the disadvantage of requiring silver nitrate to remove the protecting group, risking oxidative damage. An alternative method uses a hexamethylene S-S reagent, which is kept during the purification steps and removed by treatment with dithiothreitol (DTT). Although it is known that oxidation at the sulfur can occur, to our knowledge the reaction products were not characterized. The reactivity of the thiol allowed the formation of both intra- and intermolecular dithio bonds and was applied to the synthesis of circular oligonucleotides [9] and double stranded DNA oligomers with loops at both sides [10].

DNA synthesis is a multistep sequential procedure. For every coupling step, a high yield is required and, therefore, excess of reagents are used. Small reduction of the reaction efficiency can reduce the final yield dramatically. Therefore, careful monitoring of the synthesis steps and thorough analysis of the final compounds are necessary to avoid the formation of side products and to optimize the synthesis of oligonucleotides. Routine purification of oligonucleotides includes ion-exchange and reversed-phase liquid chromatography. Often, using these methods, closely related products are insufficiently resolved chromatographically to be spotted. In this work, we report the use of mass spectrometric

Published online August 8, 2006
Address reprint requests to Dr. J. Rozenski, Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium. E-mail: jef.rozenski@rega.kuleuven.ac.be

1044-0305/06/$32.00
doi:10.1016/j.jasms.2006.07.011
Received March 14, 2006
Revised July 12, 2006
Accepted July 12, 2006

methods for the detection and characterization of side products obtained during the synthesis of 5′ sulfhydryl containing deoxyoligonucleotides.

Materials and Methods

Oligonucleotide Synthesis

Oligonucleotide assembly was performed on an Expedite DNA synthesizer (Applied Biosystems, Foster City, CA) using the phosphoramidite approach. The standard DNA assembly protocol was used, except for 3 min coupling time using 0.07M of the newly synthesized unnatural amidites with tetrazole as the activator. Oxidations were carried out with a 0.02 M iodine solution. The oligomers were deprotected and cleaved from the solid support by treatment with concentrated aqueous ammonia:methylamine 40% 1:1 (vol/vol) (35 °C, 2 h). To this mixture, 0.05 M DTT was added for cleavage of the dithio bond and the solution was kept for 30 min at room temperature. After gel-filtration on a NAP-25 column (Sephadex G25-DNA grade; Pharmacia, Uppsala, Sweden) with water as eluent, the crude was analyzed on a Mono-Q HR 5/5 anion exchange column, after which purification was achieved on a Mono-Q HR 10/10 column (Pharmacia) with the following gradient system (A = 10 mM $NaClO_4$, 20 mM Tris buffer pH 7.4, acetonitrile 15%; B = 600 mM $NaClO_4$, 20 mM Tris buffer pH 7.4, acetonitrile 15%). The low-pressure liquid chromatography system consisted of a Merck-Hitachi L 6200 A intelligent pump (Darmstadt, Germany), a Mono Q-HR 10/10 column (Pharmacia), a Uvicord SII 2238 UV detector (Pharmacia-LKB) and a recorder. The product-containing fraction was desalted on a NAP-25 column and lyophilized.

Acrylamide Derivatization

Acrylamide is suspected to cause cancer and appropriate protecting precautions should be taken for storage, handling, and disposal of the compound. Reaction of sulfhydryl groups with acrylamide was carried out by addition of 1 μL of a 1M acrylamide solution in water and 4 μL of a 0.1M Tris buffer (pH 8.25) to 5 μL of a 50 μM oligonucleotide solution and incubation of the mixture for 2 h at 37 °C. The reaction was stopped by adding 1 μL of a 10% (vol/vol) acetic acid solution in water. For LC-MS analysis the reaction mixtures were injected directly onto the reversed-phase column.

Mass Spectrometry

Electrospray ionization mass spectra were obtained in negative ion mode on a quadrupole/time-of-flight mass spectrometer (Q-TOF-2, Micromass, Manchester, UK) equipped with a standard ionization source. The instrument was tuned to a resolution of over 8000 (full peak width at half maximum) and the mass accuracy of the instrument was less than 2 ppm using an internal calibrant (lock mass). Masses for the oligonucleotides were obtained by deconvolution of the spectra using the MaxEnt algorithm of the software (MassLynx 3.4, Micromass). Fragment ion spectra were recorded after isolation of the precursor ion in the quadrupole with a 3 u mass window and collision with argon gas in the hexapole collision cell. The collision energy was set to 80 eV, 100 eV, or 160 eV for the 11 mer, 15 mer, and 25 mer, respectively.

High-Performance Liquid Chromatography (HPLC)-Mass Spectrometry

HPLC on a C18 reversed-phase column (PepMap 0.5 × 15 mm, LC Packings, Amsterdam, The Netherlands) was performed using a buffer containing *N,N*-dimethylaminobutane (DMAB, Acros, Geel, Belgium) as ion pairing reagent and 1,1,1,3,3,3-hexafluoro-2-propanol (hexafluoro isopropanol, HFiP, Acros, Geel, Belgium) [11]. In brief, the solvent system consisted of acetonitrile 84% (vol/vol) (Fisher Scientific, Loughborough, UK) as organic phase and DMAB 0.05% (vol/vol) with HFiP 1% (vol/vol) in water as aqueous phase (pH 8.0). Oligonucleotides were eluted with a flow rate of 12 μL/min applying a gradient starting at 2% organic phase and increasing by 2% per minute during 15 min. The concentration of the oligonucleotide samples was ~50 μM and 0.2 μL (10 pmol) product was injected per run.

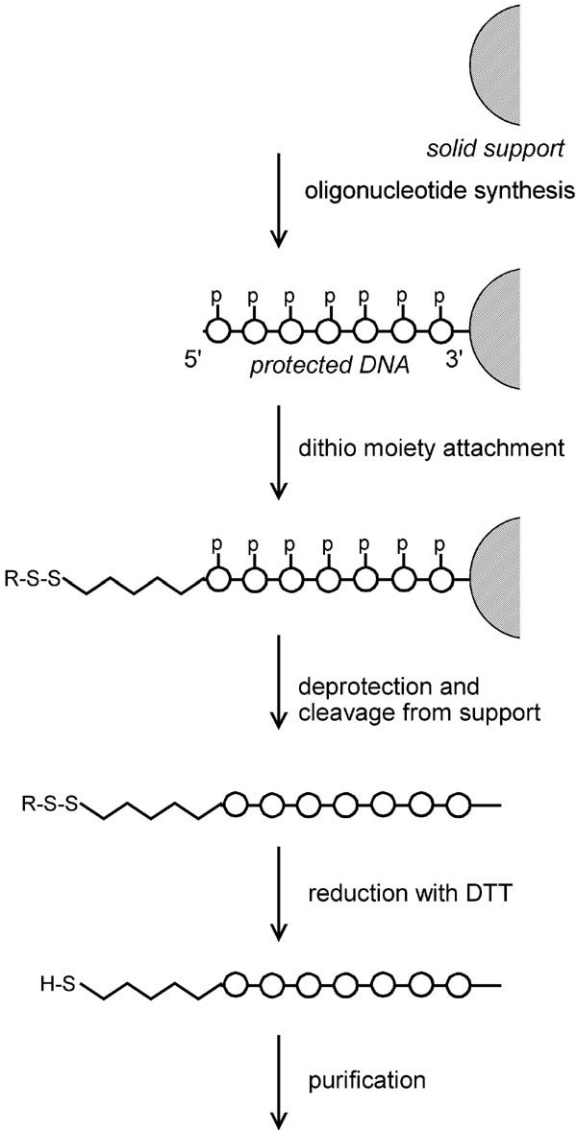
Results and Discussion

Oligonucleotide Synthesis

For the synthesis of oligonucleotides with a sulfhydryl containing moiety at the 5′ end, different approaches are possible [12, 13]. We chose a method where first a protected oligonucleotide is synthesized, followed by the attachment of the R-S-S-$(CH_2)_{11}$-group (Scheme **1**). Oxidation reactions of the intermediate phosphite triester were done with 0.02 M iodine. Finally, after cleavage from the solid support and deprotection of the oligonucleotide, the dithio bond was cleaved by treatment with DTT yielding the desired thiol containing oligonucleotide: HS-$(CH_2)_{11}$-DNA. Following this procedure, a minimum of side reactions should occur and oxidation of the sulfhydryl group by iodine is avoided. However, during LC-MS analysis several oligonucleotide species appeared to be present in the samples. The challenge was to identify and characterize the side products to give feedback and to optimize the synthesis procedure. As an example, we will describe the results for three oligonucleotides (Table 1) in further detail.

Oligonucleotide Analysis

The base peak intensity chromatograms for oligonucleotides **1–3** are shown in Figure 1. The chromatographic system separates oligonucleotides according to the



Scheme 1. Followed strategy for the synthesis of thiolated deoxyoligonucleotides.

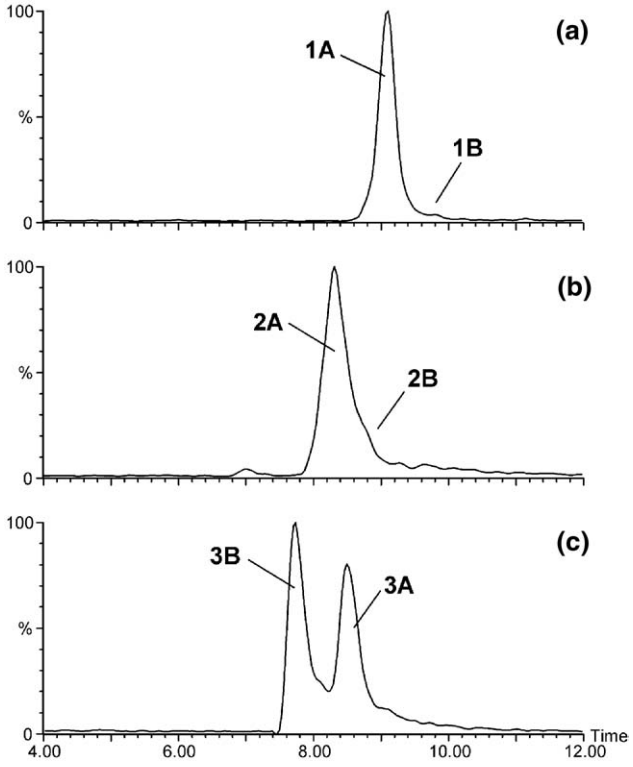


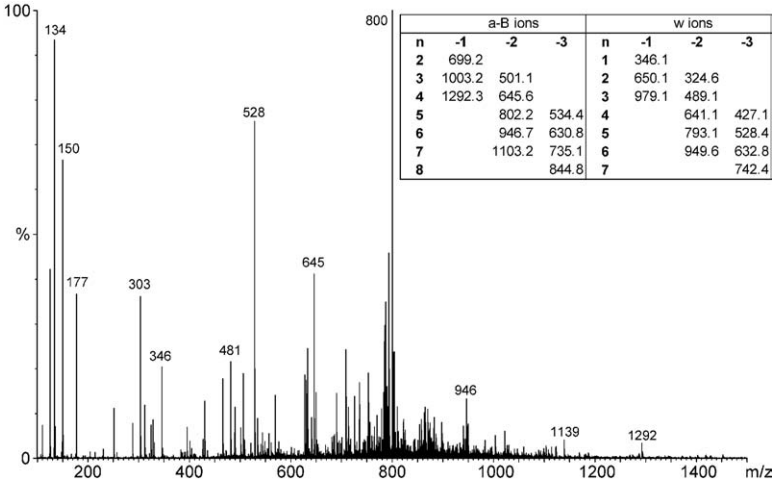
Figure 1. Base peak chromatograms for (**a**) HS-$(CH_2)_{11}$-d(T_9) (**b**) HS-$(CH_2)_{11}$-d(A_{15}), and (**c**) HS-$(CH_2)_{11}$-d(TTCACAGGTACTGGATTTGATTGTG). The peak annotations refer to structures in Table 1.

number of phosphate groups by formation of ion pairs and for unmodified oligonucleotides the retention time increases with oligonucleotide length. Here, we are dealing with oligonucleotides containing a hydrophobic aliphatic tail, and the elution order is not related to the number of phosphates, which could indicate that tertiary structures are formed during the chromatographic procedure. The major compound obtained for sequences **1** and **2** was the desired product. However, a small amount of a compound corresponding to the dithio dimer DNA-$(CH_2)_{11}$-S-S-$(CH_2)_{11}$-DNA could be detected. This product is probably formed by air oxidation after purification of the final compounds. For oligonucleotide **3**, a large amount of side product was found (about 45%) eluting earlier in the ion pair reversed-phase HPLC system with a mass 32 u higher than the expected mass. Accurate mass measurements for high molecular mass compounds do not give enough information for elemental composition determination with our mass spectrometer. In addition, the accuracy of our instruments (2 ppm) is not high enough to distinguish HS-S- and HO-(SO)- groups (mass differ-

Table 1. Synthesized oligonucleotides and their impurities identified by mass spectrometry

Sample	Peak	Oligonucleotide	Calculated mass	Found	%[b]
1	A[a]	HS-$(CH_2)_{11}$-d(T_9)	2940.6	2940.8	99
	B	d(T_9)-$(CH_2)_{11}$-S-S-$(CH_2)_{11}$-d(T_9)	5879.1	5879.5	1
2	A[a]	HS-$(CH_2)_{11}$-d(A_{15})	4900.0	4900.4	95
	B	d(A_{15})-$(CH_2)_{11}$-S-S-$(CH_2)_{11}$-d(A_{15})	9798.0	9798.6	5
3	A[a]	HS-$(CH_2)_{11}$-d(TTCACAGGTACTGGATTTGATTGTG)	7980.4	7981.0	40
	B	HO-(S=O)-$(CH_2)_{11}$-d(TTCACAGGTACTGGATTTGATTGTG)	8012.4	8012.9	60

[a]Peak shifting to an earlier retention time upon treatment with acrylamide and increasing in mass by 71 u.
[b]Approximate composition from UV_{260} data in LC-MS experiments.



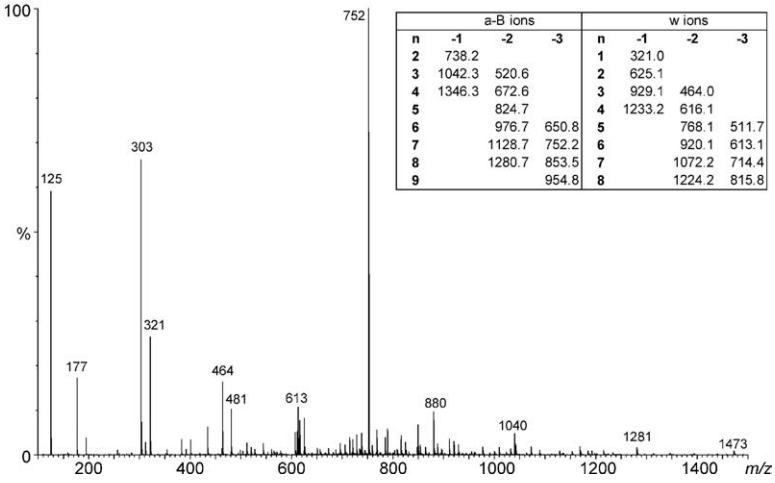
a-B ions				w ions			
n	-1	-2	-3	n	-1	-2	-3
2	699.2			1	346.1		
3	1003.2	501.1		2	650.1	324.6	
4	1292.3	645.6		3	979.1	489.1	
5		802.2	534.4	4		641.1	427.1
6		946.7	630.8	5		793.1	528.4
7		1103.2	735.1	6		949.6	632.8
8			844.8	7			742.4

Figure 2. Fragment ion spectrum of HO-(S=O)-$(CH_2)_{11}$-d(TTCACAGGTACTGGATTTGATTGTG) with assignment of the sequence ions (inset). Selected precursor 800.5 (10^-), collision energy 16 eV.

ence 18 mmu) in these oligonucleotides. Therefore, two sets of additional experiments were performed to characterize all oligonucleotides: direct sequencing by LC-MS/MS and verification of the presence of a sulfhydryl group by derivatization with acrylamide.

First, fragmentation spectra were recorded for the three samples. The fragmentation patterns from compounds **1A** and **2A** could then be used for the sequence determination of **3A** and **3B**. For this purpose, the spectra in the chromatographic peaks were summed and the mass/intensity pairs were exported to a text file that could be read by the ab initio oligonucleotide sequencing software SOS [14]. The most intense peaks in the spectrum could be assigned as the a-B or w series ions, similar as for unmodified deoxyoligonucleotides [15]. Ions at *m/z* 110, 125, 134, and 150 can be assigned for base anions of C, T, A ,and G, respectively, while single nucleotides can be found having lost a water molecule (e.g., *m/z* 303 for T). When the phosphodiester backbone is cleaved twice, the so-called internal fragments (e.g., 481 for a TT subsequence) are formed, giving information about the presence of neighbor residues [16]. In our case, the 5′ end a-B ion series gives information about the presence of the sulfhydryl group. Knowing the fragmentation patterns for compounds **1A** and **2A**, the rules could then be applied for constructs **3A** and **3B**. Compound **3A** was confirmed to be the desired product. For **3B**, the 3′ end was correct, while only a-B ions were found assuming an addition of 32 u at the 5′ end (Figure 2). For oligonucleotides **1B** and **2B**, signal intensities were too low to extract useful sequence information from their fragment ion spectra.

Derivatization with acrylamide and analysis with LC-MS(/MS) was necessary to confirm the presence of a sulfhydryl group. Acrylamide derivatization is a common and well documented technique for inactivation of cysteine residue containing peptides [17]. The reaction is a Michael type addition of acrylamide with nucleophilic compounds, in this case converting the HS-R group to NH_2-(C=O)-CH_2CH_2S-R. For oligonu-



a-B ions				w ions			
n	-1	-2	-3	n	-1	-2	-3
2	738.2			1	321.0		
3	1042.3	520.6		2	625.1		
4	1346.3	672.6		3	929.1	464.0	
5		824.7		4	1233.2	616.1	
6		976.7	650.8	5		768.1	511.7
7		1128.7	752.2	6		920.1	613.1
8		1280.7	853.5	7		1072.2	714.4
9			954.8	8		1224.2	815.8

Figure 3. fragment ion spectrum of NH_2-(C=O)-CH_2-CH_2-S-$(CH_2)_{11}$-d(T_9) with assignment of the sequence ions (inset). Selected precursor 752 (4^-), collision energy 20 eV.

cleotides **1A** and **2A**, a shift of 71 u was observed after incubation with acrylamide, accompanied by an earlier elution on the reversed-phase system. The earlier elution can be explained by the use of an ion pairing agent in the chromatographic system. The retention time is proportional with the amount of negative charges (phosphate and sulfhydryl groups). When a sulfhydryl group reacts with acrylamide, it will not be available for interaction with DMAB and the DNA:$DMAB_n$ ion pair will contain less DMAB molecules and, hence, will become more polar and provide a decreasing retention time. The fragment ion spectrum for the acrylamide adduct of **1A** along with the assignment of the peaks in the fragment ion spectrum are given in Figure 3. For both components, the fragment ion spectra confirmed the derivatization of the sulfhydryl group at the 5′ end by acrylamide. For compound **3A** the same observations were made, including a shift to an earlier retention time and a mass increase of 71 u. Compounds **1B**, **2B**, and **3B** did not react, indicating the absence of a thiol group. Unfortunately, it was not possible to obtain good quality fragment ion spectra of the derivatized sample containing **3A** and **3B** because the reaction product for **3A** coelutes with unmodified **3B**, and the masses of the precursor ions could not be selected separately in our mass spectrometer without sacrificing sensitivity. Based on the lack of reactivity with acrylamide, we can conclude that the structure for compound **3B** is most likely a an oligonucleotide with a sulfinic acid moiety HO-(S=O)-$(CH_2)_{11}$- at the 5′ end. The origin of this species remains unclear, and we speculate that it is formed by air oxidation of the sample in solution. Also, the reason why different products are formed, although all the samples were handled in the same way, could not be solved. The negative derivatization reaction excludes the presence of the isobaric HS-S-$(CH_2)_{11}$- moiety at the 5′ end. This work confirms that special precautions have to be taken during synthesis of thiolated DNA oligomers to avoid the formation of oxidized side products.

Conclusions

Following the synthesis of thiol containing DNA, side products were found in different samples. The sequence of the correct compounds was proven by recording fragment ion spectra. For the modification described in this paper, the fragmentation follows the same rules as for unmodified DNA. Reaction with acrylamide showed (1) that the thiol group is intact for the correct compounds and from the fragment ion spectra we could see that the addition of acrylamide proceeds at the 5′ thiol moiety and (2) that the side products did not have thiol groups available for modification. From the molecular mass of the side products, we identified these in the case of samples **1** and **2** as dithio-linked dimers. In sample **3**, the major component was a product with 32 u higher than expected. Here, the oligonucleotide was characterized as the sulfinic acid oxidation product. In all cases, the side products were oxidation products, presumably originating from air oxidation. To our knowledge, the reaction products have not been characterized. For amino acids and proteins on the contrary, the oxidation of cysteine and methionine is well known [18]. In future, extra precaution will be taken to avoid the formation of such oxidation products.

Acknowledgments

The authors thank G. Schepers for technical assistance and the Research Foundation-Flanders (FWO-Vlaanderen) for financial support through grants 1.5.019.6 (JR) and G.0603.06 (AVA).

References

1. Okamoto, T.; Suzuki, T.; Yamamoto, N. Microarray Fabrication with Covalent Attachment of DNA Using Bubble Jet Technology. *Nature Biotech.* **2000,** *18,* 438–441.
2. Yu, F.; Yao, D.; Knoll, W. Oligonucleotide Hybridization Studied by a Surface Plasmon Diffraction Sensor (SPDS). *Nucleic Acids Res.* **2004,** *32e75.*
3. Fixe, F.; Dufva, M.; Telleman, P.; Christensen, C. B. V. One-Step Immobilization of Aminated and Thiolated DNA Onto Poly(Methylmethacrylate) (PMMA) Substrates. *Lab. Chip.* **2004,** *4,* 191–195.
4. Tuukanen, S.; Virtanen, J.; Hytönen, V. P.; Kulomaa, M. S.; Törma, P. Fabrication of DNA Monolayers on Gold Substrates and Guiding of DNA with Electric Field. *Rev. Adv. Mater. Sci.* **2003,** *5,* 228–231.
5. Demers, L. M.; Ginger, D. S.; Park, S.-J.; Li, Z.; Chung, S.-W.; Mirkindagger, C. A. Direct Patterning of Modified Oligonucleotides on Metals and Insulators by Dip-Pen Nanolithography. *Science* **2002,** *296,* 1836–1838.
6. Sproat, B. S.; Beijer, B.; Rider, P.; Neuner, P. The Synthesis of Protected 5′Mercapto-2′,5′-Dideoxyribonucleoside-3′-O-Phosphoramidites; Uses of 5′-Mercapto-Oligoribonucleosides. *Nucleic Acids Res.* **1987,** *15,* 4837–4848.
7. Connolly, B. A.; Rider, P. Chemical Synthesis of Oligonucleotide Containing a Free Sulphydryl Group and Subsequent Attachment of Thiol Specific Probes. *Nucleic Acids Res.* **1985,** *13,* 4485–4502.
8. Gaur, R. K.; Sharma, P.; Gupta, K. C. A Simple Method for the Introduction of Thiol Group at 5′-Termini of Oligodeoxynucleotides. *Nucleic Acids Res.* **1989,** *17,* 4404.
9. Gao, H.; Yang, M.; Patel, R.; Cook, A. F. Circularization of Oligonucleotides by Disulfide Bridge Formation. *Nucleic Acids Res.* **1995,** *23,* 2025–2029.
10. Gao, H.; Yang, M.; Patel, R.; Cook, A. F. Stabilization of Double-Stranded Oligonucleotides Using Backbone-Linked Disulfide Bridges. *Nucleic Acids Res.* **1995,** *23,* 285–292.
11. Apffel, A.; Chakel, J. A.; Fischer, S.; Lichtenwalter, K.; Hancock, W. S. Analysis of Oligonucleotides by HPLC-Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1997,** *69,* 1320–1325.
12. Sinha, N. D.; Cook, R. M. The Preparation and Application of Functionalized Synthetic Oligonucleotides. III. Use of H-Phosphonate Derivatives of Protected Amino-Hexanol and Mercapto-Propanol or -Hexanol. *Nucleic Acids Res.* **1988,** *16,* 1659–1669.
13. Kumar, A.; Advani, S.; Dawar, H.; Talwar, G. P. A Simple Method for Introducing a Thiol Group at the 5′-End of Synthetic Oligonucleotides. *Nucleic Acids Res.* **1991,** *19,* 4561.
14. Rozenski, J; McCloskey, J. A. SOS: A Simple Interactive Program for ab Initio Oligonucleotide Sequencing by Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2002,** *13,* 200–203.
15. McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply-Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992,** *3,* 60–70.
16. Rozenski, J.; McCloskey, J. A. Determination of Nearest Neighbors in Nucleic Acids by Mass Spectrometry. *Anal. Chem.* **1999,** *71,* 1454–1459.
17. Garzitto, M.; Rovatto, L.; Hamdan, M. On-Line Liquid Chromatography/Electrospray Tandem Mass Spectrometry to Investigate Acrylamide Adducts with Cysteine Residues: Implications for Polyacrylamide Gel Electrophoresis Separations of Proteins. *Rapid Commun. Mass Spectrom.* **1998,** *12,* 484–488.
18. Houde, D.; Kauppinen, P.; Mhatre, R.; Lyubarskaya, Y. Determination of Protein Oxidation by Mass Spectrometry and Method Transfer to Quality Control. *J. Chromatorgr.* **2006,** *1123,* 189–198.